PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Unexplained Mortality During the U.S. COVID-19 Pandemic: Retrospective Analysis of Death Certificate Data and Critical Assessment of Excess Death Calculations
AUTHORS	Fairman, Kathleen; Goodlet, Kellie; Rucker, James; Zawadzki, Roy

VERSION 1 – REVIEW

REVIEWER	Vieira, André Universidade Nova de Lisboa, Escola Nacional de Saúde Pública
REVIEW RETURNED	14-Apr-2021

GENERAL COMMENTS	This is an interesting study, aiming to understand the contribution of the excess deaths reported beyond the top 8 natural causes of death in the USA, verified in the pandemic period. However, I state some suggestions and commentaries, hoping they can improve your manuscript. The use of different benchmark periods may lead to generate data beyond the necessary for the scope of this work. I believe the average of the 5 seasons would be fine. However, if you believe the 3 periods are important for the analysis, its relevance should be clear in the text. I believe the table 1 could be improved for some better understanding and intuitive view. For example, there isn't a "total explained deaths" column, and Change could also be address as "Excess Deaths". Also, the figure should be better explained, regarding the weeks and seasons. For example, some vertical lines dividing into pandemic awareness (January 19 through March 28); initial pandemic peak (March 29 through May 30); and pandemic post-peak (May 31 through September 26). The proportion of Total Explained deaths / Total deaths doesn't change so much across all seasons in study. In fact, in a rapid view, I find that 2019/2020 had lower rates of unexplained deaths/total deaths than the previous years. This should be carefully addressed, because although there is a higher number of deaths, it seems that the proportions of the explained vs unexplained deaths relating to the total of deaths didn't differ, and it may be expected. The problem of lacking some registering for psychiatric-cause deaths still exists, but it is not so clear if it changed in COVID-19. NEC definition in this study was performed by their authors or is the definition from ICD-10? If so, please indicate the reference. What are the top 8 main natural causes? The appendix presents 10 categories… Are the "R" CODES in or out of the NEC? Is there any explanation for the fact that categories F00 to F99 from the

ICD are not included in the NHCS? This should be clearer in the text.

The authors should consider also other potential causes for these unexplained deaths beyond the psychiatric-cause deaths. In fact, it may be plausible, but External causes of morbidity and mortality (https://icd.who.int/browse10/2019/en#/XX) may overlap the psychiatric-cause deaths.

1.

I have some doubts about what the research question is. In lines 141 to 144 ("The primary research question was whether and to what extent interpretation of the excess-death metric is potentially vulnerable to unexplained mortality. A secondary research question was when unexplained deaths occurred relative to specific phases of the pandemic") the objectives don't refer to COVID-19 mortality, but the results presented separated that. Also, what is considered "potentially vulnerable"? It sounds a little confusing to state it clear.

2.

I have some doubts if "interpretability" is the right word in line 46 ("To assess interpretability of the excess-death metric, we parsed excess deaths in the coronavirus-2019 (COVID-19) pandemic into explained versus unexplained causes."), and that compromises the Objectives definition of the study.

The results section starts with "Of 287,957-306,267 excess all-cause deaths through September 2020". These two numbers are unclear in the following of the abstract; we should expect three numbers for each season comparison.

In the section of "Setting", states of EUA should appear.

The relevance of the missing classification of "psychiatric-cause deaths" in EUA coding appears too late, and probably would benefit from some comparisons with countries with data where this classification is performed, even in periods previous COVID-19. In lines 54-55 of abstract, this information appears "natural causes, and respiratory causes including COVID-19—with no category for psychiatric-cause deaths" but not along the methods or introduction of the main text.

3.

I struggled to understand that seasons 2, 3 and 5 were related to the last seasons starting from 2019/2020. I really think it would be easier to understand if only one benchmark more representative would be selected. Probably some modeling like ARIMA would fit this purpose better than descriptive statistics, considering the 6-season period. Still, authors should clarify the objectives to understand the relevance of COVID-19 specific assessment or if they care solely about the unexplained vs explained deaths metrics.

5.

There is no statement about ethics, even it is clear the use of data from publicly databases.

6.

I believe that this outcome should be also better explained: the underlying cause of by itself could be the specific causes of deaths, and that was not what the results showed. The results showed Explained Reported UCOD, COVID-19, Explained Reported UCOD, not COVID-19, UCOD Was NEC and

Unreported: All-Cause Death, No UCOD. Also, some clarification of the difference between NEC and Unreported: All-Cause Death, No UCOD should improve the understanding.

7.
I feel this article needs more detailing about the statistics. The sentence "Analyses were performed using open-source analytic tools" should be replaced for more accurate information. As said before, probably a statistical analysis model that uses time series data probably would be more appropriate for this goal.
There is no information at all about what kind of descriptive or analytic statistics were applied (as means, etc.).

9.
This is in line with the information of the research question.

10.
The tables and figures are good for interpretation but should be reviewed considering the suggestions above.

11.
Interpretation of data with other countries (regarding other explanations for the unexplained excess deaths in EUA) should be performed, even in periods before the COVID-19.

12.
I think some limitations should also be addressed to the lack of control about the coding processes, perhaps even more during the pandemics. I believe there could exist much more insights for this problematic also. For example, it should be of reflection the fact that the % of unexplained deaths / total deaths decreased in 2019 / 2020, perhaps because the coding of deaths was improved because of COVID-19 (?), or some people that died with COVID-19 would be classified with other unknown death cause if there was no pandemic?

I apologize if I had misunderstood some of the study information.

REVIEWER	Ballotari, Paola ATS Val Padana
REVIEW RETURNED	02-May-2021

GENERAL COMMENTS	I agree with the Authors, the pandemic period excess deaths was not due only to Covid 19 or related-causes, as the respiratory diseases. Our unpublished study covering the first wave confirms the excess deaths for all causes, especially infectious diseases (also but not only sepsis), mental and behavioural disorders, and diseases of the nervous system. Disentangle whether was due to misclassification or to virus indirect effects, is very difficult. Moreover, it would be important to distinguish between deaths with Covid 19 (or related illness) as underlying cause and deaths with positive test (occurred within 30 days or at any time?). I also agree that NCHS classification sound unsuitable for analyse direct and indirect effect of Covid 19 by means of underlying causes of deaths. I don't understand if NCHS classification it's provisional or not, please explain Methods-results

	The comparison would be easier, using only one benchmark (e.g. 3-season). Perhaps, it would be interesting to explains differences between the three benchmarks (the latter looks different than the first two). It would be interesting to add stratification analysis by sex and age classes. The NEC causes strongly increased between the fortieth and fiftieth season week, could you provide some explanatory hypotesis?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. André Vieira, Universidade Nova de Lisboa
Comments to the Author:
This is an interesting study, aiming to understand the contribution of the excess deaths reported beyond the top 8 natural causes of death in the USA, verified in the pandemic period.
However, I state some suggestions and commentaries, hoping they can improve your manuscript.

The use of different benchmark periods may lead to generate data beyond the necessary for the scope of this work. I believe the average of the 5 seasons would be fine. However, if you believe the 3 periods are important for the analysis, its relevance should be clear in the text.
Authors' response: Both reviewers expressed concern about the presentation of multiple benchmarks in the data table. We believe that the use of multiple benchmarks provides an important sensitivity analytic check on results but agree that the data presentation in Table 1 was confusing. We retained the multiple benchmarks in the Figure, deleted the prior 2- and 3-season benchmarks from Table 1 (page 26), and moved the 2- and 3-season benchmarks to a new Appendix 2. We explained the decision to use multiple benchmarks on page 9, lines 292-295).

I believe the table 1 could be improved for some better understanding and intuitive view. For example, there isn't a "total explained deaths" column, and Change could also be address as "Excess Deaths".
Authors' response: We concur with this observation. We added a column for Total Explained, bolded the excess deaths to make them easier for readers to find, and labeled the percentage row as "% of excess deaths." (Note that excess deaths are, by definition, all-cause deaths, so none of the specific mortality categories shown in the column headers can appropriately be labeled as "excess deaths."). Please see Table 1 on page 26.

Also, the figure should be better explained, regarding the weeks and seasons. For example, some vertical lines dividing into pandemic awareness (January 19 through March 28); initial pandemic peak (March 29 through May 30); and pandemic post-peak (May 31 through September 26).
Authors' response: We have added dividing lines to Figure 1 and thank the reviewer for the suggestion.

The proportion of Total Explained deaths / Total deaths doesn't change so much across all seasons in study. In fact, in a rapid view, I find that 2019/2020 had lower rates of unexplained deaths/total deaths than the previous years. This should be carefully addressed, because although there is a higher number of deaths, it seems that the proportions of the explained vs unexplained deaths relating to the

total of deaths didn't differ, and it may be expected. The problem of lacking some registering for psychiatric-cause deaths still exists, but it is not so clear if it changed in COVID-19.

Authors' response: We appreciate the question because it pointed us to an aspect of the method that needed clarification. The excess-death calculation does not represent the count of total deaths; it represents <u>change</u> in total deaths compared with historical experience. For that reason, the denominator, excess deaths, is total death minus historical experience death (i.e., change in all-cause deaths), and the proportions represent the percentage of <u>change</u> in total death that is explained versus unexplained by <u>changes</u> in reported specific causes of death. We have added an explanatory sentence to the final paragraph of the Introduction to explain this point (page 7, lines 218-221). We have also edited the Table 1 row labels to make them address this point more clearly (page 26). We also include more explicit definitions of changes in the Methods section (page 10, lines 306-312 and lines 315-318). Finally, we added text to clarify that the proportion unexplained decreased during the height of the pandemic, compared with the other periods (Abstract, lines 79-80; Results page 11, lines 344 and 350).

NEC definition in this study was performed by their authors or is the definition from ICD-10? If so, please indicate the reference.

Authors' response: The term "not elsewhere classified" is included in the ICD-10 definition for the R00-R99 category. This information appears in Supplementary Appendix 1, but we have added a detailed explanation to the Methods in an expanded section on the data source, and we now cite the ICD-10 (page 8, lines 252-255).

What are the top 8 main natural causes? The appendix presents 10 categories… Are the "R" CODES in or out of the NEC?

Authors' response: We have added much more information about the contents of the files to the expanded section on the data source in the Methods section (page 8, lines 247-261).

In addition to the top 8 causes of natural death, the files include COVID-19, other respiratory conditions, septicemia, and the NEC category, for a total of 12. Previously, the Supplementary Appendix 1 failed to list COVID-19 and showed only 11 total categories; now, all 12 categories are listed with their corresponding ICD-10 code ranges. This change should make the file structure much clearer for readers.

Is there any explanation for the fact that categories F00 to F99 from the ICD are not included in the NHCS? This should be clearer in the text.

Authors' response: We are unable to answer the question of why the NCHS did not include psychiatric or injurious causes in the weekly mortality files that it made available in May 2020 for pandemic-period analysis. We have not found any published information on this question and did not wish to speculate about it. We now explicitly state that the data do not include specific psychiatric or injurious causes (page 8, lines 256-261). In the Discussion section, we also note that in March 2021, the NCHS released total monthly counts of injurious causes of death, including overdose and suicide, for the United States as a whole, reported through August 2020. These were the subject of an analysis published in May 2021, which we also cite and discuss (page 13, lines 205-209 and page 14, line 456 to page 15, line 471).

The authors should consider also other potential causes for these unexplained deaths beyond the psychiatric-cause deaths. In fact, it may be plausible, but External causes of morbidity and mortality (https://icd.who.int/browse10/2019/en#/XX) may overlap the psychiatric-cause deaths.

Authors' response: We agree with the reviewer that it is important to mention the possibility that injurious causes, including suicides and overdose, could have been reported as underlying causes of death instead of the psychiatric conditions underlying those injurious deaths. We now mention absence of those codes from the available data (page 8, lines 259-261) and include in the Discussion a description of the need for information about these

5

causes (page 14, lines 451-456). We also added a Limitations bullet noting that we are not able to address the question of either injurious or psychiatric causes using the currently available data.

1.

I have some doubts about what the research question is. In lines 141 to 144 ("The primary research question was whether and to what extent interpretation of the excess-death metric is potentially vulnerable to unexplained mortality. A secondary research question was when unexplained deaths occurred relative to specific phases of the pandemic") the objectives don't refer to COVID-19 mortality, but the results presented separated that. Also, what is considered "potentially vulnerable"? It sounds a little confusing to state it clear.

Authors' response: We have edited the text for clarity and changed the language to that used earlier (line 191) when introducing the topic: "degree to which the utility and interpretation of excess death calculations are potentially compromised by unreported or unspecified causes of mortality" (page 7, lines 217-223).

2.

I have some doubts if "interpretability" is the right word in line 46 ("To assess interpretability of the excess-death metric, we parsed excess deaths in the coronavirus-2019 (COVID-19) pandemic into explained versus unexplained causes."), and that compromises the Objectives definition of the study.

Authors' response: We removed the phrase, as it was confusing, and the previous sentence regarding actionable policy information makes the issue clear (Abstract, lines 47-48).

The results section starts with "Of 287,957-306,267 excess all-cause deaths through September 2020". These two numbers are unclear in the following of the abstract; we should expect three numbers for each season comparison.

Authors' response: The numbers represent a range, depending on the seasonal benchmark. We clarified the language (Abstract, line 62).

In the section of "Setting", states of EUA should appear.

Authors' response: We are not sure we understand the comment, but the setting in the United States was mentioned under Design (Abstract, line 51). The phrase "U.S. states" would equate to "United States states." We apologize if we misunderstood the reviewer.

The relevance of the missing classification of "psychiatric-cause deaths" in EUA coding appears too late, and probably would benefit from some comparisons with countries with data where this classification is performed, even in periods previous COVID-19. In lines 54-55 of abstract, this information appears "natural causes, and respiratory causes including COVID-19—with no category for psychiatric-cause deaths" but not along the methods or introduction of the main text.

Authors' response: We have added to the Methods section more detail about the absence of psychiatric and injurious causes from the files (page 8, lines 256-261). We removed this phrase from the Abstract because it was confusing, and to comply with word count limits. Although investigation of international variation in certification of all causes of death would be a massive undertaking and out of scope of the present investigation, we did add more detail on international differences in COVID-19 death certification to the first paragraph of the Introduction (page 5, lines 144-148).

3.

I struggled to understand that seasons 2, 3 and 5 were related to the last seasons starting from 2019/2020. I really think it would be easier to understand if only one benchmark more representative would be selected. Probably some modeling like ARIMA would fit this purpose better than descriptive statistics, considering the 6-season period. Still, authors should clarify the objectives to understand

the relevance of COVID-19 specific assessment or if they care solely about the unexplained vs explained deaths metrics.

Authors' response: We believe that edits made in response to this reviewer's other comments will make the purpose clearer. We did not perform ARIMA because we found that descriptive analyses produced the same results as more sophisticated modeling techniques, and because the descriptive analyses permitted us to parse excess death (changes in all-cause deaths) into explained and unexplained proportions. We now explain this decision in Data Analyses, a new subheader in the Methods section (page 9, lines 283-287), and provide a comparison of our work and previously published results in the first paragraph of the Discussion section (page 12, lines 382-385).

5.

There is no statement about ethics, even it is clear the use of data from publicly databases.

Authors' response: We have added an ethical statement (page 17, lines 545-548). If the editors would prefer a different format, please advise us.

6.

I believe that this outcome should be also better explained: the underlying cause of by itself could be the specific causes of deaths, and that was not what the results showed. The results showed Explained Reported UCOD, COVID-19, Explained Reported UCOD, not COVID-19, UCOD Was NEC and Unreported: All-Cause Death, No UCOD. Also, some clarification of the difference between NEC and Unreported: All-Cause Death, No UCOD should improve the understanding.

Authors' response: We believe that several changes recommended elsewhere by this reviewer, including added detail about the data file, the explanation of the ICD-10 NEC category, the addition of a Total Explained column to Table 1, and the clarified definition of excess death as representing a change over historical experience, addressed this concern.

7.

I feel this article needs more detailing about the statistics. The sentence "Analyses were performed using open-source analytic tools" should be replaced for more accurate information. As said before, probably a statistical analysis model that uses time series data probably would be more appropriate for this goal.

Authors' response: We agree that more detail about the analytic procedures was appropriate and have added it (page 8, lines 275-276, lines 283-287, lines 289-290, lines 295-296). With respect to ARIMA or modeling, please see our previous response to this question.

There is no information at all about what kind of descriptive or analytic statistics were applied (as means, etc.).

Authors' response: This information was in the Methods section, but we expanded on it so that readers do not miss it (page 9, lines 289-292).

9.

This is in line with the information of the research question.

10.

The tables and figures are good for interpretation but should be reviewed considering the suggestions above.

11.

Interpretation of data with other countries (regarding other explanations for the unexplained excess deaths in EUA) should be performed, even in periods before the COVID-19.

Authors' response: We have added more detail about international COVID-19 CODA practices to the Introduction (page 5, lines 144-148), and on geographic disparity in underlying risk

factors to the Discussion (page 15, lines 477-486).

12.
I think some limitations should also be addressed to the lack of control about the coding processes, perhaps even more during the pandemics. I believe there could exist much more insights for this problematic also. For example, it should be of reflection the fact that the % of unexplained deaths / total deaths decreased in 2019 / 2020, perhaps because the coding of deaths was improved because of COVID-19 (?), or some people that died with COVID-19 would be classified with other unknown death
cause if there was no pandemic?
Authors' response: We believe that the changes made in response to other comments addressed this concern. As we noted above in response to those comments, cross-national variations in coding practices are described in more detail in the first paragraph of the Introduction, and the focus on change in explained versus unexplained causes of death is explained in the final paragraph of the Introduction.

I apologize if I had misunderstood some of the study information.
Authors' response: We greatly appreciate your helpful observations.
Kind Regards

Reviewer: 2
Dr. Paola Ballotari, ATS Val Padana
Comments to the Author:
I agree with the Authors, the pandemic period excess deaths was not due only to Covid 19 or related-causes, as the respiratory diseases. Our unpublished study covering the first wave confirms the excess deaths for all causes, especially infectious diseases (also but not only sepsis), mental and behavioural disorders, and diseases of the nervous system. Disentangle whether was due to misclassification or to virus indirect effects, is very difficult. Moreover, it would be important to distinguish between deaths with Covid 19 (or related illness) as underlying cause and deaths with positive test (occurred within 30 days or at any time?). I also agree that NCHS classification sound unsuitable for analyse direct and indirect effect of Covid 19 by means of underlying causes of deaths.
Authors' response: Thank you for these observations. In response to a different reviewer's comment, we added information about wide cross-national variations in COVID-19 death-certification practices to the Introduction (page 5, lines 144-148).

I don't understand if NCHS classification it's provisional or not, please explain
Authors' response: Thank you for pointing out this omission. We have added this information to the Methods section under Design and Data Source (page 7, lines 228-231).

Methods-results
The comparison would be easier, using only one benchmark (e.g. 3-season). Perhaps, it would be interesting to explains differences between the three benchmarks (the latter looks different than the first two).
Authors' response: The other reviewer also suggested that the comparison would be more easily understood with just one benchmark, although that reviewer suggested 5 years rather than 3. We believe that the use of multiple benchmarks provides an important sensitivity analytic check on results but agree that the data presentation in Table 1 was confusing. We retained the multiple benchmarks in the Figure, deleted the prior 2- and 3-season benchmarks from Table 1 (page 26), and moved the 2- and 3-season benchmarks to a new Appendix 2. We explained the decision to use multiple benchmarks on page 9, lines 292-295). We also indicate the range of all-cause death totals across the 3 benchmark periods in the first paragraph of the Results section (page 10, lines 323-324).

It would be interesting to add stratification analysis by sex and age classes.
Authors' response: We agree, but unfortunately the currently available files that include cause-of-death categories do not include age or sex. We now include in the Discussion section a description of populations that might be particularly vulnerable to psychiatric or injurious causes of death, and mention that availability of the full, final cause-of-death files for 2020 will facilitate research of the type you suggest (page 13, lines 412-419; page 14, lines 445-447).

The NEC causes strongly increased between the fortieth and fiftieth season week, could you provide some explanatory hypotesis?
Authors' response: We now point out this change in the Results (page 10, lines 324-325) and offer two possible explanations for it in the Discussion (page 12, lines 393-399).

VERSION 2 – REVIEW

REVIEWER	Vieira, André
	Universidade Nova de Lisboa, Escola Nacional de Saúde Pública
REVIEW RETURNED	29-Jul-2021

GENERAL COMMENTS	Dear authors,
	Thank you for your responses and modification. I truly believe the manuscript has improved and it´s much easier to understand the data.
	1.The changes performed in table 1 are much better and comprehensible now (to me). I don´t know the sense of it, but shouldn´t be also the benchmarks values below the "Total changes, pandemic awareness through end of observation" ?
	2.I believe, table 1, in the subtitles, I suggest something like f 2019-20 values minus average benchmark values. Instead of f 2019-20 values minus benchmark values. and I think the "% of excess deaths" should also have some description under the table, as it refers to excess deaths in the season 2019-2020. Also, I understand the table already have a lot of information, and the relevance of the "% of excess deaths". But, regarding I understand the focus on the unexplained vs explained excess deaths, I miss some comparison with the "regularity" unexplained deaths proportion. I mean: Is there any difference between seasons in the % of Total Unexplained/Total Deaths? I think that information is not there and helps to understand if there was more or less % attributed to unexplained deaths in the pandemic season. For example, in the Initial pandemic peak period, this fraction was 30,4% for the benchmark period and 27,1% in 2019-2020. In fact, I believe this relation have decreased in all seasons for 2019-2020. I think this simple analysis is also very important to the purpose of your study, as it is not mandatory to be stated in this table.
	3.As you presented the results in the main table with a benchmark of 5 years, I suggest to give some focus on that and state the other benchmark analysis of (2 and 3 years) as secondary or as

| | complement. It is only a suggestion, as you also have some data of it in figure 1.

4.In figure 1, I believe giving some description of what color represents each end of weeks could help clarifying the image.

5.In this sentence "(…) the files include four additional categories: miscellaneous respiratory conditions, (e.g., nasopharyngitis, sinusitis, pneumothorax); septicemia; COVID-19; and nonspecific-cause deaths.13,14 These deaths, described in ICD-10 nomenclature as "symptoms, signs, and abnormal clinical and laboratory findings not elsewhere classified" (NEC; ICD-10 range R00-R99), include "ill-defined and unknown cause of mortality" (R99),13,14 a code commonly used pending forensic investigation of injurious death.23"
"These deaths" refer to nonspecific-cause deaths? I believe so, but please, clarify better, because the sentence is after different causes of death.

6."In the section of "Setting", states of EUA should appear." - I apologize for my mistake here, it was a wrong word. I only wanted to say that it should made reference to the country in that section, the USA, but you are right as it appears before, that´s ok. Thank you for the effort.

7.Finally, I still have some concern about the second research question. I feel that it probably would need some other study design to be explore it, as it is stated. I believe the question you raised is more an interpretation of the results of the first research question than a research question by itself of this study. I kindly ask you to give your view about this.
Thanks for your work. |
|---|---|

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Dr. André Vieira, Universidade Nova de Lisboa
Comments to the Author:
Dear authors,

Thank you for your responses and modification. I truly believe the manuscript has improved and it´s much easier to understand the data. **We appreciate the reviewer's observations and contributions to the improvements in the manuscript.**

1.The changes performed in table 1 are much better and comprehensible now (to me). I don´t know the sense of it, but shouldn´t be also the benchmarks values below the
"Total changes, pandemic awareness through end of observation"? **We appreciate the suggestion, agree, and added the two rows.**

2.I believe, table 1, in the subtitles, I suggest something like
 f 2019-20 values minus average benchmark values. **We added the word "mean" to the description (see Table 1).**
Instead of
f 2019-20 values minus benchmark values.
and I think the "% of excess deaths" should also have some description under the table, as it refers to

excess deaths in the season 2019-2020. **We agree and added the description: "Change from mean benchmark count to 2019-20, expressed as a percentage of total excess all-cause deaths in 2019-20" (see newly added text in footnote to Table 1).**

Also, I understand the table already have a lot of information, and the relevance of the "% of excess deaths". But, regarding I understand the focus on the unexplained vs explained excess deaths, I miss some comparison with the "regularity" unexplained deaths proportion. I mean:

Is there any difference between seasons in the % of Total Unexplained/Total Deaths? I think that information is not there and helps to understand if there was more or less % attributed to unexplained deaths in the pandemic season.

For example, in the Initial pandemic peak period, this fraction was 30,4% for the benchmark period and 27,1% in 2019-2020. In fact, I believe this relation have decreased in all seasons for 2019-2020. I think this simple analysis is also very important to the purpose of your study, as it is not mandatory to be stated in this table. **We remain concerned that the analysis requested by the reviewer does not address the purpose of the research, which was to assess the percentage of <u>excess</u> death that was unexplained. Nonetheless, we realize that other journal readers may have the same question. Accordingly, we added a Supplementary Appendix (2), which shows the percentage of total deaths (not excess deaths) that were unexplained.**

The added analysis showed that during the entire study period (from awareness through end of observation), the count of unexplained deaths as a percentage of total deaths was 30.6% using the 5-year mean benchmark and 29.8% in 2019-20, indicating a slight 0.8 percentage-point decline in 2019-20. Comparing 2019-20 with the 5-season benchmark, the unexplained percentage was 1.0 percentage point higher in the pandemic awareness period (30.8% versus 29.8%, respectively), 3.3 percentage points lower in the pandemic peak period (27.1% vs. 30.4%), and approximately equal (0.3 percentage points lower, 30.8% vs. 31.1%) in the pandemic post-peak period.

As shown in Appendix 2, and consistent with the absolute death count changes presented in Table 1, these minimal changes in the unexplained portion of total deaths occurred because both explained and unexplained deaths increased during the pandemic, resulting in a substantial contribution of unexplained death to excess death.

3.As you presented the results in the main table with a benchmark of 5 years, I suggest to give some focus on that and state the other benchmark analysis of (2 and 3 years) as secondary or as complement. It is only a suggestion, as you also have some data of it in figure 1. **As the reviewer correctly surmised, our text on page 10 relies not only on Table 1, which shows the numeric detail for the 5-year benchmark, but also on Figure 1. For that reason, we did not make additional changes to the text.**

4.In figure 1, I believe giving some description of what color represents each end of weeks could help clarifying the image. **We believe that the information requested by the reviewer is in the Figure legend (page 27), but we added a specific explanation of the lines associated with each of the seasonal benchmarks. If we have misunderstood, we apologize.**

5.In this sentence "(…) the files include four additional categories: miscellaneous respiratory conditions, (e.g., nasopharyngitis, sinusitis, pneumothorax); septicemia; COVID-19; and nonspecific-cause deaths.13,14 These deaths, described in ICD-10 nomenclature as "symptoms, signs, and abnormal clinical and laboratory findings not elsewhere classified" (NEC; ICD-10 range R00-R99), include "ill-defined and unknown cause of mortality" (R99),13,14 a code commonly used pending forensic investigation of injurious death.23"

"These deaths" refer to nonspecific-cause deaths? I believe so, but please, clarify better, because the sentence is after different causes of death. **Thank you for pointing out this problem. We have corrected "These deaths" to "Nonspecific cause deaths" on page 8.**

6."In the section of "Setting", states of EUA should appear." - I apologize for my mistake here, it was a wrong word. I only wanted to say that it should made reference to the country in that section, the USA, but you are right as it appears before, that´s ok. Thank you for the effort.

7.Finally, I still have some concern about the second research question. I feel that it probably would need some other study design to be explore it, as it is stated. I believe the question you raised is more an interpretation of the results of the first research question than a research question by itself of this study. I kindly ask you to give your view about this. **We agree with the reviewer, as we describe in the Discussion, that a complete investigation of the causes of unexplained deaths would indeed require a different design, which in turn would require more and better detailed data on causes of death. Nonetheless, we originally formulated the secondary research question as an exploratory analysis and cannot make a retroactive change to what we did at that time. In researching the response to the reviewer's question, we realized that one sentence in the Methods section, regarding when the analyses presented in the paper were planned, was misleading. We removed that sentence (page 10), moved that information to the end of the Introduction (page 7), and reworded the language on page 7 to prevent confusion.**
Thanks for your work.